AD-756 844

A MAXIMUM-LIKELIHOOD MULTIPLE-HYPOTHESIS TESTING ALGORITHM, WITH AN APPLICATION TO MONOPULSE DATA EDITING

E. J. Kelly

Massachusetts Institute of Technology

Prepared for:

Federal Aviation Administration
Air Force Systems Command

9 February 1973

AD756844

Technical Note 1973-7

E. J. Kelly

A Maximum-Likelihood Multiple-Hypothesis Testing Algorithm, with an Application to Monopulse Data Editing

9 February 1973

Prepared for Federal Aviation Administration by

Lincoln Laboratory

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

LEXINGTON, MASSACHUSETTS

Availability is unlimited. Document may be released to the National Technical Information Service, Springfield, Virginia 22151, for sale to the public.

28

1. Introduction

There are many data processing problems in which decisions must be made, based on noisy data, which do not fit the well-known pattern of a binary test between two hypotheses. Instead, there are N distinct hypotheses ($N > 2$), each of which may be composite (i.e., contain internal parameters), and an algorithm is sought whereby each observation, or sample, implies a unique choice of one of the hypotheses. We are not interested in randomized tests or sequential tests, and the discussion is carried out in the simple context of a finite - dimensional sample space with hypotheses expressed in terms of well-behaved probability densities, each containing a finite number of real parameters.

In this note, we suggest a simple algorithm for multiple-hypothesis testing, based on the maximum-likelihood technique for deciding between hypothesis pairs. The algorithm is not optimum in any sense, but has the virtue that it works, while possessing considerable intuitive appeal. The procedure arose out of the consideration of a particularly simple, but practical, problem which requires a multiple-hypothesis formulation, namely the detection of interference in a monopulse direction-finding system. This problem, which arises in Air Traffic Control surveillance, is discussed here as an illustration of the testing algorithm.

2. Multiple-Hypothesis Testing

Let x represent a sample, i.e., a point in a multidimensional observation space, from which a decision must be made among N hypotheses, H_i ($i = 1, \ldots, N$). The decision rule must be a decomposition of the observation space, X, into N disjoint sets, D_i ($i = 1 \ldots, N$), so that a sample falling into D_i implies the choice of hypothesis i, while

$$\bigcup_{i=1}^{N} D_i = X$$

Hypothesis H_i is characterized by a probability density,

$$f_i(x \mid \alpha_i),$$

which contains a finite set of parameters, α_i. The i^{th} parameter set, α_i, is a point in a finite dimensional parameter space, A_i.

If we had only to decide between hypotheses H_i and H_j, we would follow the generalized maximum likelihood principle, maximizing each probability density over its respective parameters and comparing their ratio to a threshold. In other words, we would accept H_i over H_j whenever

$$\frac{\sup\limits_{\alpha_i \in A_i} f_i(x \mid \alpha_i)}{\sup\limits_{\alpha_j \in A_j} f_i(x \mid \alpha_j)} \geq \lambda_{ij} \tag{1}$$

In terms of the log likelihood functions

$$\ell_i(x) \equiv \log \sup_{\alpha_i \in A_i} f_i(x \mid \alpha_i) , \tag{2}$$

we write

$$T(i/j): \ \ell_i(x) - \ell_j(x) \geq \log \lambda_{ij} \tag{3}$$

This expression is to be read: "testing H_i over H_j, accept H_i whenever the indicated inequality is true". If we set the threshold, λ_{ij}, equal to unity, we have $\log \lambda_{ij} = 0$, and the test amounts to selecting the "most likely explanation" of the data. In general we do not use $\lambda_{ij} = 1$ because we anticipate the need to control the inevitable decision-making errors, often in an unsymmetrical way between the two hypotheses. Expression (3) can be rewritten in a suggestive way by introducing two threshold parameters, μ_i and μ_j, as follows:

$$T\,(i\,/\,j):\ \ell_i(x) - \mu_i \geq \ell_j(x) - \mu_j \tag{4}$$

In words: "H_i is accepted if $\ell_i(x)$ exceeds its threshold, μ_i, by at least as much as the amount by which $\ell_j(x)$ exceeds its threshold, μ_j". With just two hypothesis , only the difference,

$$\mu_i - \mu_j = \log \lambda_{ij} \ ,$$

is relevant, but (4) can be generalized to the N - hypothesis case quite easily. A threshold, μ_i, is associated with each hypothesis and the "excess", $\ell_i(x) - \mu_i$, is computed; the hypothesis with the largest excess is accepted. A precise definition, in which ambiguities are resolved, is as follows.

For a given sample let

$$M(x) = \operatorname*{Max}_{j=1}^{N} \left[\ell_j(x) - \mu_j \right] \tag{5a}$$

Then,

$$M(x) = \ell_k(x) - \mu_k \tag{5b}$$

for at least one value of k. We assign x to the set D_k, where k is the smallest index for which (5b) is true.

This algorithm has N-1 free constants, say

$$\mu_j - \mu_1 \quad , \quad j = 2, \cdots, N,$$

which can be chosen (in principle) to control N-1 decision errors. These errors are expressed in terms of the probability of choosing H_i when H_j is true with $\alpha_j \in B_j$, a subset of A_j.

In this algorithm, H_i is chosen only when every other H_j is rejected according to a test of the type (3). It might be thought that a more general algorithm could be developed by introducing N (N-1) / 2 constants, λ_{ij} $(i > j)$, and the corresponding number of regions, R_{ij}, defined by

$$\ell_i(x) - \ell_j(x) \geq \log \lambda_{ij} \qquad (i > j) \quad . \tag{6}$$

Then each point, x, is either in R_{ij} or its complement, for each distinct pair, (i, j). If $x \in R_{ij}$, then H_i is "preferred" over H_j, and hence for each point, x, all the pairwise "preferences" are established by the definitions (6). The difficulty is that there is no guarantee that one hypothesis will be preferred over all others since the transitivity of the preference relation is not an automatic consequence of (6). This is most easily seen for N = 3 and can be illustrated in a two-dimensional sample space as shown in Figure 1.

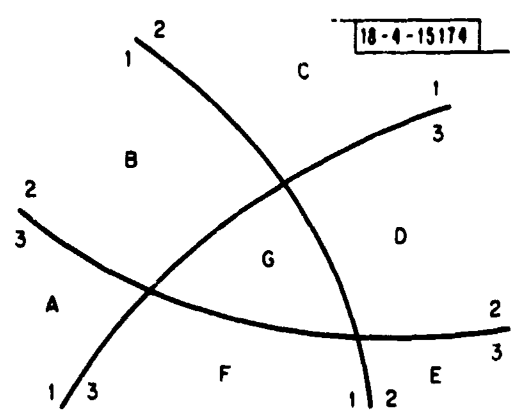

Fig. 1. Preference regions.

Each line in Figure 1 represents a boundary defining a region R_{ij} and its complement for one pair of the three hypotheses H_1, H_2 and H_3. The numbers indicate the preferred hypothesis on either side of the line.

In this algorithm, H_i is chosen only when every other H_j is rejected according to a test of the type (3). It might be thought that a more general algorithm could be developed by introducing N (N-1) / 2 constants, λ_{ij} $(i > j)$, and the corresponding number of regions, R_{ij}, defined by

$$\ell_i(x) - \ell_j(x) \geq \log \lambda_{ij} \qquad (i > j) \quad . \tag{6}$$

Then each point, x, is either in R_{ij} or its complement, for each distinct pair, (i, j). If $x \in R_{ij}$, then H_i is "preferred" over H_j, and hence for each point, x, all the pairwise "preferences" are established by the definitions (6). The difficulty is that there is no guarantee that one hypothesis will be preferred over all others since the transitivity of the preference relation is not an automatic consequence of (6). This is most easily seen for N = 3 and can be illustrated in a two-dimensional sample space as shown in Figure 1.

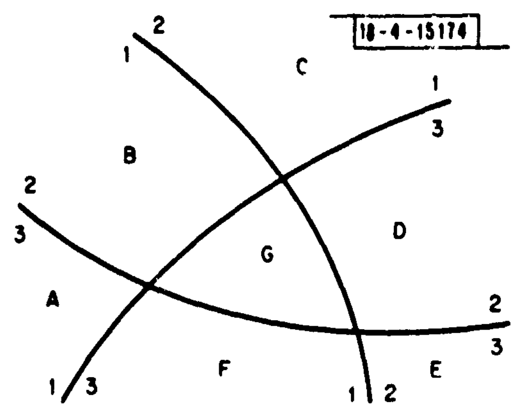

Fig. 1. Preference regions.

Each line in Figure 1 represents a boundary defining a region R_{ij} and its complement for one pair of the three hypotheses H_1, H_2 and H_3. The numbers indicate the preferred hypothesis on either side of the line.

In each of the six regions, A, B, C, D, E and F the pairwise preferences establish one of the 3! possible hierarchies of choice, each with a clear "first choice". Thus, in region A, H_1 is preferred to H_3 and H_3 is preferred to H_2. However, in region G, H_1 is preferred to H_2, H_2 is preferred to H_3, and H_3 is preferred to H_1, hence the preferences are inconsistent with no clear choice. If, in Figure 1, we interchange the numbers 2 and 3 on the H_2-H_3 boundary, then G becomes a region of consistent preferences while C and F are not.

This situation will not happen in our algorithm since a clear choice is always made. The boundary between D_1 and D_2 is the surface defined by:

$$\ell_1(x) - \mu_1 = \ell_2(x) - \mu_2 ,$$

while the D_2 - D_3 boundary is described by

$$\ell_2(x) - \mu_2 = \ell_3(x) - \mu_3$$

The intersection of these two boundaries is a subspace in which

$$\ell_1(x) - \mu_1 = \ell_3(x) - \mu_3 ,$$

which is contained in the D_1 -D_3 boundary. Thus regions like G in Figure 1 do not arise.

3. The Monopulse Data Editing Problem

For our purpose, an "amplitude-comparison monopulse system" can be modeled as an antenna-receiver system in which rf signals are derived from each of two effective antennas having coincident phase centers. It does not matter whether the two "antennas" are realized by a pair of horn-terminated waveguides facing a single reflector, or by a pair of feed networks connected to the elements of an array, so long as the phase centers coincide and the main-beam voltage gains are real (i.e. negligible phase shift across either main beam). The two rf signals are amplified and demodulated to produce two in-phase and two quadrature components of video, each containing additive

Gaussian noise. These four noise processes are mutually independent with identical spectra. In the problem at hand, the receiver has detected and synchronized its own timing on the early portion (preamble) of an anticipated incident signal. The waveform is simple binary pulse-amplitude-modulation (PAM OOK). The four video waveforms are sampled once at each bit position, and we are concerned with deciding among various hypotheses regarding the true state of affairs on the basis of one of these four-dimensional samples. A separate decision is made for each bit position in the waveform (A more general problem, not treated here, concerns multiple-hypothesis testing based on a sequence of samples as one decision.)

Let the in-phase and quadrature samples from antenna 1 be combined as the real and imaginary parts of a complex observable, Z_1, and let the corresponding components from antenna 2 be combined to form Z_2 The pair (Z_1, Z_2) represents our basic sample, and hypotheses will be formulated as probability densities in these two complex variables. The sample (Z_1, Z_2) always contains the receiver noise components (n_1, n_2), which are independent, zero-mean, complex Gaussian variables, satisfying

$$E\, n_i^2 = 0 \text{ and } E\, |n_i|^2 = 2\, \sigma_a^2 \qquad i = 1, 2 \qquad (7)$$

The noise variance, σ_a^2, is fixed by the receiver characteristics, and is considered known.

A signal, coming from a single direction (azimuth), in the main beam of both antennas, will produce sample components, $(A_1 e^{i\psi}, A_2 e^{i\psi})$, of the same phase. The phase, ψ, is essentially the rf phase of the incident wave at the common phase center of the antennas, and the signal amplitude, A_i, is the product of the incident wave amplitude and the voltage gain of the i^{th} beam in the signal direction. Signal direction is determined [1] from an estimate of A_1/A_2, but this aspect of the processing is not of direct interest here. If there are two signals in the main beam, arriving from different directions, then each will contribute sample components of the form just discussed, but each signal has a different rf phase, and the two signals have unequal amplitude ratios, so that the resultant sample components must be modelled as $(A_1 e^{i\psi_1}, A_2 e^{i\psi_2})$, where A_1 and A_2 are arbitrary

positive constants, and ψ_1 and ψ_2 are arbitrary angle variables. The second signal is interpreted either as interference from another source, or as off-azimuth multipath from the main signal. Obviously, the same model represents the case of three or more main-beam signals.

A final possibility is a combination of one or more signals arriving from outside the main antenna beams, where the antenna gains are complex and vary rapidly with angle. Such a combination, representing side and/or back lobe interference has been modelled as Gaussian noise, characterized by sample components (m_1, m_2) which are statistically identical to the receiver noise components, except for their variance

$$E\,|m_i|^2 = 2\,\sigma_b^2 \qquad i = 1,\,2 \tag{8}$$

When this interference is present the total noise components of the sample are still independent, zero-mean complex Gaussian variables, characterized by relations like (7), with σ_a^2 replaced by $\sigma_a^2 + \sigma_b^2$, considered known (this assumption is discussed below).

With these models we can distinguish six possible hypotheses, depending upon the number of main-beam signals and the presence or absence of side/back lobe interference. The hypotheses are defined in the following table.

Number of main-beam signals	Side / Back Lobe Interference	Hypothesis
0	NO	1
1	NO	2
≥ 2	NO	3
0	YES	4
1	YES	5
≥ 2	YES	6

Table I

If no main-beam signals are judged to be present, it is assumed that the signal information bit is a "zero". If one or more main-beam signals are found, this information bit is assumed to be a "one". Interference is reported if it appears to be present as side/back lobe interference, or if more than one main-lobe signal appears. The presence of interference of either kind is used to aid the message decoding algorithm, and also to inhibit direction-finding on the signal (by means of the amplitude ratio), since a spurious value would quite likely be found. Thus, the output of the decision algorithm is an estimated information bit and an interference flag ("one" is present), according to Table II.

Hypothesis	Information Bit	Interference Flag
1	0	0
2	1	0
3	1	1
4	0	1
5	1	1
6	1	1

Table II

Note that hypotheses H_3, H_5 and H_6, although statistically distinct, all lead to the same response. Moveover, it will turn out that H_3 and H_6 are indistinguishable from the data, and H_6 will later be dropped. H_3 and H_5 must be tested separately, even though the resulting decision regions, D_3 and D_5, are combined to determine system response. The term "data editing", as used here, refers to the detection of interference and the resulting use of the interference flag in decoding and direction finding. These latter topics are not discussed here, and we return to the specific formulation of the multiple-hypothesis testing problem.

According to our models, the probability densities for the six hypotheses have the forms

$$f_i(Z_1, Z_2|\alpha_i) = (2\pi\sigma_a^2)^{-2} \exp\left[-\frac{L_i^2(Z_1, Z_2|\alpha_i)}{2\sigma_a^2}\right] \quad i = 1, 2,$$

and (9)

$$f_i(Z_1, Z_2|\alpha_i) = \left[2\pi(\sigma_a^2 + \sigma_b^2)\right]^{-2} \exp\left[-\frac{L_i^2(Z_1, Z_2|\alpha_i)}{2(\sigma_a^2 + \sigma_b^2)}\right] i = 4,$$

where

$$L_1^2(Z_1, Z_2) = L_4^2(Z_1, Z_2) = |Z_1|^2 + |Z_2|^2 \quad \text{(no parameters)} \tag{10}$$

$$L_2^2(Z_1, Z_2|A_1, A_2, \psi) = L_5^2(Z_1, Z_2|A_1, A_2, \psi) = |Z_1 - A_1 e^{i\psi}|^2 + |Z_2 - A_2 e^{i\psi}|^2$$

$$L_3^2(Z_1, Z_2|A_1, A_2, \psi_1, \psi_2) = L_6^2(Z_1, Z_2|A_1, A_2, \psi_1, \psi_2) = |Z_1 - A_1 e^{i\psi_1}|^2 + |Z_2 - A_2 e^{i\psi_2}|^2$$

Since the variable parameters in all cases are internal to the functions L_i^2, the required maxima of the probability densities involve the minima of the L_i^2. Since L_1^2 and L_4^2 involve no parameters, there is no minimization to perform, while L_3^2 and L_6^2 can be made equal to zero by the parameter choices

$$A_i e^{i\psi_i} = Z_i \quad i = 1, 2$$

The remaining expression is

$$L_2^2 = L_5^2 = |Z_1|^2 + |Z_2|^2 - 2A_1 \mathrm{Re}\,(e^{-i\psi}Z_1) - 2A_2\,\mathrm{Re}\,(e^{-i\psi}Z_2) + A_1^2 + A_2^2$$

$$= |Z_1|^2 + |Z_2|^2 + \left[A_1 - \mathrm{Re}\,(e^{-i\psi}Z_1)\right]^2 + \left[A_2 - \mathrm{Re}\,(e^{-i\psi}Z_2)\right]^2$$

$$- \left[\mathrm{Re}\,(e^{-i\psi}Z_1)\right]^2 - \left[\mathrm{Re}\,(e^{-i\psi}Z_2)\right]^2 \quad .$$

From

$$\cos^2\theta = \tfrac{1}{2} + \tfrac{1}{2}\cos^2\theta$$

we infer that

$$\left[\mathrm{Re}\,(e^{-i\psi}Z_1)\right]^2 = \tfrac{1}{2}\,|e^{-i\psi}Z_1|^2 + \tfrac{1}{2}\mathrm{Re}\,(e^{-i\psi}\,Z_1)^2$$

$$= \tfrac{1}{2}\,|Z_1|^2 + \tfrac{1}{2}\,\mathrm{Re}\,(e^{-2i\psi}\,Z_1^2) \quad .$$

Thus

$$L_2^2 = L_5^2 = \tfrac{1}{2}(|Z_1|^2 + |Z_2|^2) + \left[A_1 - \mathrm{Re}\,(e^{-i\psi}Z_1)\right]^2 + \left[A_2 - \mathrm{Re}\,(e^{-i\psi}Z_2)\right]^2$$

$$- \tfrac{1}{2}\,\mathrm{Re}\,\left[e^{-2i\psi}(Z_1^2 + Z_2^2)\right] \quad .$$

This expression is clearly minimized by the choice

$$\hat{\psi} = \tfrac{1}{2}\arg(Z_1^2 + Z_2^2) \quad , \qquad (11)$$

so that

$$\mathrm{Re}\,\left[e^{-2i\hat{\psi}}(Z_1^2 + Z_2^2)\right] = |Z_1^2 + Z_2^2| \quad ,$$

together with the choices

$$\hat{A}_1 = \mathrm{Re}\,(e^{-i\hat{\psi}}Z_1)$$
$$\hat{A}_2 = \mathrm{Re}\,(e^{-i\hat{\psi}}Z_2) \qquad (12)$$

The resulting minimum is

$$\inf_{\alpha_i \in A_i} L_2^2 = \inf_{\alpha_i \in A_i} L_5^2 = \tfrac{1}{2}\left(|Z_1|^2 + |Z_2|^2\right) - \tfrac{1}{2}\left|Z_1^2 + Z_2^2\right|$$

We introduce the notation

$$\begin{aligned} P &\equiv |Z_1|^2 + |Z_2|^2 \\ Q &\equiv \left|Z_1^2 + Z_2^2\right| , \end{aligned} \tag{13}$$

and summarize our results in the form of log likelihood ratios:

$$\begin{aligned}
\ell_1(Z_1, Z_2) &= -2 \log (2\pi\,\sigma_a^2) - \frac{1}{2\,\sigma_a^2}\, P \\
\ell_2(Z_1, Z_2) &= -2 \log (2\pi\sigma_a^2) - \frac{1}{4\sigma_a^2}\,(P - Q) \\
\ell_3(Z_1, Z_2) &= -2 \log (2\pi\sigma_a^2) \\
\ell_4(Z_1, Z_2) &= -2 \log \left[2\pi\,(\sigma_a^2 + \sigma_b^2)\right] - \frac{1}{2\,(\sigma_a^2 + \sigma_b^2)}\, P \\
\ell_5(Z_1, Z_2) &= -2 \log \left[2\pi\,(\sigma_a^2 + \sigma_b^2)\right] - \frac{1}{4(\sigma_a^2 + \sigma_b^2)}\,(P - Q) \\
\ell_6(Z_1, Z_2) &= -2 \log \left[2\pi\,(\sigma_a^2 + \sigma_b^2)\right]
\end{aligned} \tag{14}$$

4. The Decision Regions

According to equations (5) in Section 2, the testing algorithm involves the differences $\ell_i(x) - \mu_i$, where the μ_i are arbitrary constants to be assigned later. Thus, the constant terms in equations (14) can be absorbed into the μ_i. In addition, the $\ell_i(x)$ can all be multiplied by a fixed constant without changing the decision regions. We therefore ignore the log - terms in (14) and multiply by the factor $(-2\sigma_a^2)$. The effect of the minus sign is to change from Max to Min in equations (5), hence the decision regions are based on

$$M(Z_1, Z_2) = \operatorname*{Min}_{j=1}^{6} \left[g_j(Z_1, Z_2) + \mu_j \right] \tag{15}$$

where

$$\begin{aligned} g_1(Z_1, Z_2) &= P \\ g_2(Z_1, Z_2) &= \tfrac{1}{2}(P - Q) \\ g_3(Z_1, Z_2) &= 0 \\ g_4(Z_1, Z_2) &= \frac{1}{R} P \\ g_5(Z_1, Z_2) &= \frac{1}{2R}(P - Q) \\ g_6(Z_1, Z_2) &= 0 \end{aligned} \tag{16}$$

and

$$R \equiv \frac{\sigma_a^2 + \sigma_b^2}{\sigma_a^2} > 1 \quad , \tag{17}$$

The μ_i in (15) are new arbitrary constants. We can simplify things by assigning those sample points for which

$$g_i(Z_1, Z_2) + \mu_i < g_k(Z_1, Z_2) + \mu_k \quad , \quad k \neq i \tag{18}$$

to region D_i, and making an arbitrary assignment of points for which two or more of the $g_i + \mu_i$ are equal. These boundary points will not contribute to any integrals expressing decision probabilities or errors, since the functions g_k are all continuous.

We note that the sign of $\mu_3 - \mu_6$ is independent of the data and hence one of the two hypotheses, H_3 or H_6, is always preferred over the other, depending on the choice of the μ_i. This simply means that the data cannot support a decision between H_3 and H_6, and H_6 is now dropped from our discussion with the understanding that H_3, as represented by $g_3(Z_1, Z_2) + \mu_3$, represents the composite case of two or more main - beam signals, with or without side/back lobe interference.

The decision region for H_i is simply the intersection of the regions defined by (18) for all values of k distinct from i. Since the $g_k(Z_1, Z_2)$ depend only on the quantities P and Q, these two statistics are sufficient for decision between all five hypotheses. Moreover, the regions defined by (18), expressed in the (P, Q) - plane, are all half - planes, bounded by various straight lines. The defining conditions are as follows, in terms of the coordinates P and Q.

H_1 is chosen if

$$\begin{aligned} P + Q &< 2(\mu_2 - \mu_1) , \\ P &< \mu_3 - \mu_1 , \\ P &< \frac{R}{R-1}(\mu_4 - \mu_1) , \text{ and} \\ (2R-1)P + Q &< 2R(\mu_5 - \mu_1). \end{aligned} \tag{19a}$$

H_2 is chosen if

$$
\begin{aligned}
P + Q &\geq 2(\mu_2 - \mu_1), \\
P - Q &< 2(\mu_3 - \mu_2), \\
(R-2)P - RQ &< 2R(\mu_4 - \mu_2), \text{ and} \\
P - Q &< \frac{2R}{R-1}(\mu_5 - \mu_2).
\end{aligned}
\tag{19b}
$$

H_3 is chosen if

$$
\begin{aligned}
P &\geq \mu_3 - \mu_1, \\
P - Q &\geq 2(\mu_3 - \mu_2), \\
P &\geq R(\mu_3 - \mu_4), \quad \text{and} \\
P - Q &\geq 2R(\mu_3 - \mu_5).
\end{aligned}
\tag{19c}
$$

H_4 is chosen if

$$
\begin{aligned}
P &\geq \frac{R}{R-1}(\mu_4 - \mu_1), \\
(R-2)P - RQ &\geq 2R(\mu_4 - \mu_2), \\
P &< R(\mu_3 - \mu_4), \text{ and} \\
P + Q &< 2R(\mu_5 - \mu_4).
\end{aligned}
\tag{19d}
$$

H_5 is chosen if

$$(2R - 1)\,P + Q \geq 2R\,(\mu_5 - \mu_1)\,,$$

$$P - Q \geq \frac{2R}{R-1}\,(\mu_5 - \mu_2)\,,$$

$$P - Q < 2R\,(\mu_3 - \mu_5), \text{ and} \tag{19e}$$

$$P + Q \geq 2R\,(\mu_5 - \mu_4)\,.$$

Arbitrary assignments of boundary points have been made in equations (19) which involve ten distinct straight lines, corresponding to the ten pairs of hypotheses. P and Q are inherently positive, and $P > Q$, by the Pythagorean inequality, hence only that portion of the first quadrant in the (P, Q) - plane between the P - axis and the line $P = Q$ is attainable from the data. Unless a boundary line crosses that portion, it can have no effect on any decision region. When parallel lines enter into the definition of a decision region, only one will be effective, depending on the relative values of the μ_i. Thus, a considerable variety of shapes is available for the regions D_i, and error probabilities will have to be formulated and assigned in order to choose among them.

Note that H_2 is chosen over H_1 (refer to Table I) if $P + Q$ exceeds a constant. This is the detection statistic obtained by Hofstetter and DeLong[1] in their analysis of amplitude - comparison monopulse. Details may be found in that paper concerning the estimation of signal direction from the parameter estimates given in (11) and (12), once H_2 has been accepted. We see also that H_2 is chosen over the signal - plus - interference hypotheses, H_3 and H_5, if $P - Q$ is sufficiently small. This test, which is related to a requirement that the monopulse beam outputs be in phase, has been obtained by McAulay[2] for H_2 against H_5, and DeLong[3] for H_2 against H_3. An approximation to the P-Q test has also been obtained by McGarty[4].

The boundary lines and the hypothesis pairs they separate are the following:

(H_1, H_2)	$P + Q = 2(\mu_2 - \mu_1)$	(20a)
(H_4, H_5)	$P + Q = 2R(\mu_5 - \mu_4)$	(20b)
(H_2, H_3)	$P - Q = 2(\mu_3 - \mu_2)$	(20c)
(H_2, H_5)	$P - Q = \frac{2R}{R-1}(\mu_5 - \mu_2)$	(20d)
(H_3, H_5)	$P - Q = 2R(\mu_3 - \mu_5)$	(20e)
(H_1, H_3)	$P = \mu_3 - \mu_1$	(20f)
(H_1, H_4)	$P = \frac{R}{R-1}(\mu_4 - \mu_1)$	(20g)
(H_3, H_4)	$P = R(\mu_3 - \mu_4)$	(20h)
(H_1, H_5)	$(2R - 1)P + Q = 2R(\mu_5 - \mu_1)$	(20i)
(H_2, H_4)	$(R - 2)P - RQ = 2R(\mu_4 - \mu_2)$	(20j)

In practice, R will be relatively large compared to unity, and in this case line (20i) is nearly parallel to lines (20f), (20g) and (20h). Also, line (20j) becomes nearly parallel to lines (20c), (20d) and (20e). In this limit, our testing regions become insensitive to the assumption that σ_b^2, the interference power, is known. An interesting possibility would be to consider $\sigma_a^2 + \sigma_b^2$ to be another unknown parameter in hypotheses 4 and 5. Returning to equation (14), we would find that the estimates of this parameter are P/4 (on H_4) and (P -Q)/4 (on H_5), while g_4 and g_5 would be changed from the expressions given in (16) to

$$g_4(Z_1, Z_2) = 4\sigma_a^2 \log P$$

$$g_5(Z_1, Z_2) = 4\sigma_a^2 \log(P - Q) \tag{21}$$

The (P, Q) - plane still suffices to define the decision regions, but some of the boundary lines would no longer be straight.

Returning to equations (20), we note that H_2, H_3 and H_5 are separated from one another entirely by lines of the form $P - Q =$ constant, while H_1, H_3 and H_4 are separated among themselves by the value of P alone. The presence of a single main - beam signal is detected by the value of $P + Q$, whether the background noise is receiver noise or random interference.

If two boundary lines intersect, and also have a hypothesis in common, such as (20a) and (20d) which both involve H_2, then a third boundary line must also pass through the intersection, in this case (20i), separating H_1 and H_5 (this can be verified by direct substitution.) This is an example the phenomenon discussed at the end of Section 2, and many other three - line intersections may be anticipated. Only eight of the ten possible cases actually arise (corresponding to the ten possible hypothesis triplets), however, because of the parallelism of many of our boundary lines.

There are many possibilities for the actual shapes of our decision regions, depending upon the choice of the μ_i. In order to give all the boundary lines a chance to traverse the attainable sample space, the right sides of equations (20a) through (20i) must all be positive, since the corresponding left sides have that property. This results in the inequalities

$$\mu_3 > \mu_5 > \mu_2 > \mu_1 \quad , \qquad \text{and}$$

$$\mu_5 > \mu_4 > \mu_1 \quad , \tag{22}$$

which establishes an ordering between all pairs except μ_2 and μ_4. We must also order the three sets of parallel lines. Thinking of R as large compared to unity, we assume further that

$$R(\mu_5 - \mu_4) > \mu_2 - \mu_1 \quad , \tag{23a}$$

$$R(\mu_3 - \mu_5) > \mu_3 - \mu_2 > \frac{R}{R-1}(\mu_5 - \mu_2), \text{ and} \tag{23b}$$

$$R(\mu_3 - \mu_4) > \mu_3 - \mu_1 > \frac{R}{R-1}(\mu_4 - \mu_1). \tag{23c}$$

These conditions are not strictly necessary, but are eminently reasonable, considering the relations (22).

In Figure 2 we give an illustration, consistent with inequalities (22) and (23), showing all ten lines and all eight triple intersections. The letters labelling the lines refer to equations (20). In Figure 3 we show the resulting decision regions, obtained by application of equations (19). There are, in effect, four adjustable constants in the choice of the regions D_i. With the assumptions of equations (23), they may be taken to be the (P, Q) - coordinates of the point where D_1, D_2 and D_4 meet, the distance from there to the point where D_2, D_4 and D_5 meet, and the distance from this latter point to the point where D_3, D_4 and D_5 meet.

In order to choose these constants we must assign four error probabilities. There is no guarantee that any set of four such probabilities can be attained, but reasonable compromises can probably be found by trial and error. It should be recalled that in our problem H_3 and H_5 result in the same system response, hence D_3 and D_5 should be united and thought of as a single decision region. One possible choice of error probabilities, each of which should be "small", would be E(2/1), E(3 + 5/1), E(3 + 5/4) and E(3 + 5/2), where E(i + j/k) stands for the probability of declaring H_i or H_j to be true when H_k is actually valid. E(2/1) is the simple false - alarm probability of declaring an information "one" to be present (but interference absent) when only receiver noise is contained in the data. In the same true situation, E(3 + 5/1) is the probability of falsely declaring the presence of interference along with an information "one". E(3 + 5/4), which probably cannot be made as small as E(2/1), to which it is analogous, is the probability of correctly recognizing the presence of side/back lobe interference, but falsely declaring an information "one". A bound must be assigned to this error for a range of values of σ_b^2, or R. The last error, E(3 + 5/2), is the

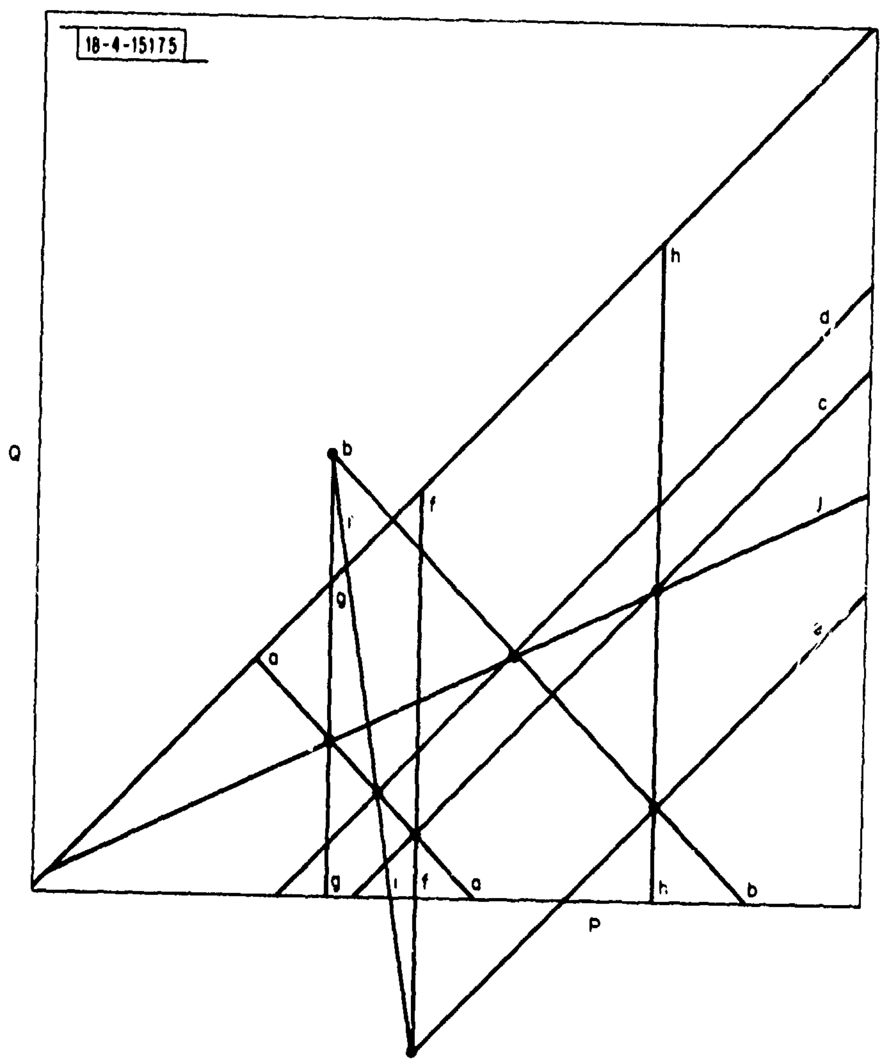

Fig. 2. Decision boundaries.

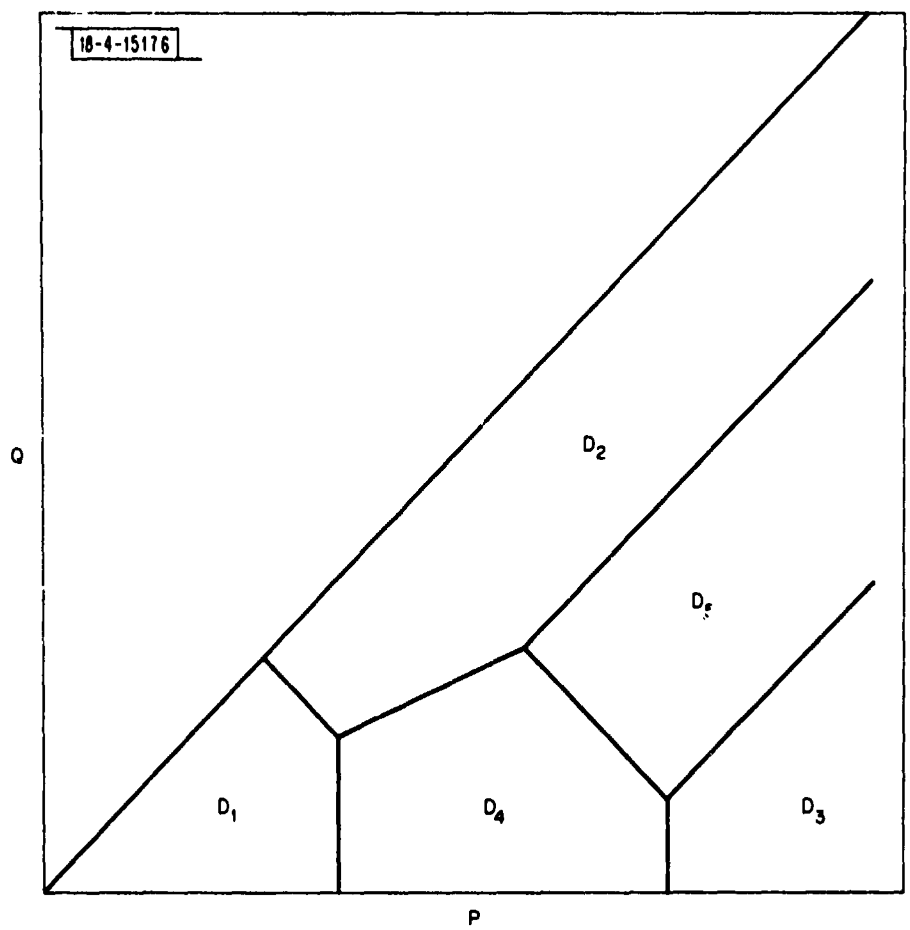

Fig. 3. Decision regions.

probability of correctly recognizing the presence of an information "one", but falsely flagging the presence of interference, hence casting doubt on the decoding and rejecting the measurement of signal direction. In general, this probability wil will depend on the signal - to - noise ratio which exists under H_2, and E(3 + 5/2) will be subjected to a bound over a range of signal - to - noise ratios. The nature of these errors is summarized in Table III.

True Parameters		Error	Reported Parameters	
Information	Interference	Type	Information	Interference
0	0	E(2/1)	1	0
0	0	E(3 + 5/1)	1	1
0	1	E(3 + 5/4)	1	1
1	0	E(3 + 5/2)	1	1

Table III

The error probabilities are not easily computed. On hypotheses H_1 and H_4 the probability density in the (Z_1, Z_2) sample space has the form

$$f(Z_1, Z_2) = (2\pi \sigma^2)^{-2} \exp - (P/2\sigma^2) , \tag{24}$$

where, according to (9), $\sigma^2 = \sigma_a^2$ on H_1 and $\sigma^2 = \sigma_a^2 + \sigma_b^2$ on H_4. From (24) it follows that the marginal probability density of the statistic P is

$$f(P) = \frac{P}{4\sigma^4} \exp - (P/2\sigma^2) . \tag{25}$$

However the joint marginal of P and Q appears to be difficult to obtain, and the regions of integration required are awkward to work with.

In the remaining case, E(3 + 5/2), a useful approximate evaluation can be made, as follows. We have

$$Z_1 = A_1 e^{i\psi} + n_1 \ , \quad Z_2 = A_2 e^{i\psi} + n_2$$

and hence

$$\begin{aligned} P &= |A_1 e^{i\psi} + n_1|^2 + |A_2 e^{i\psi} + n_2|^2 \\ &= A_1^2 + A_2^2 + 2\mathrm{Re}\left[e^{-i\psi}(A_1 n_1 + A_2 n_2)\right] + |n_1|^2 + |n_2|^2 \end{aligned} \tag{26}$$

and

$$Q = |(A_1 e^{i\psi} + n_1)^2 + (A_2 e^{i\psi} + n_2)^2| \ . \tag{27}$$

In the expression for Q we write

$$\begin{aligned} Q &= |(A_1 + e^{-i\psi} n_1)^2 + (A_2 + e^{-i\psi} n_2)^2| \\ &= |A_1^2 + A_2^2 + 2e^{-i\psi}(A_1 n_1 + A_2 n_2) + e^{-2i\psi}(n_1^2 + n_2^2)| \\ &= (A_1^2 + A_2^2)\,|1 + Z| \ , \end{aligned}$$

where

$$Z \equiv 2e^{-i\psi} \frac{A_1 n_1 + A_2 n_2}{A_1^2 + A_2^2} + e^{-2i\psi} \frac{n_1^2 + n_2^2}{A_1^2 + A_2^2} \ .$$

We put

$$Z = X + iY$$

and expand:

$$|1 + Z| = 1 + X + \tfrac{1}{2} Y^2 + \ldots..$$

to second order in $|Z|$, assumed small compared to unity. This will be the case for large signal - to - noise ratio, hence

$$Q = A_1^2 + A_2^2 + 2\,\mathrm{Re}\left[e^{-i\psi}(A_1 n_1 + A_2 n_2)\right] + \mathrm{Re}\left[e^{-2i\psi}(n_1^2 + n_2^2)\right] + \frac{2}{A_1^2 + A_2^2}\left\{\mathrm{Im}\left[e^{-i\psi}(A_1 n_1 + A_2 n_2)\right]\right\}^2$$

Thus P and Q agree through first order, and we can write

$$P - Q = |e^{-i\psi} n_1|^2 + |e^{-i\psi} n_2|^2 - \mathrm{Re}\left[e^{-2i\psi}(n_1^2 + n_2^2)\right] - \frac{2}{A_1^2 + A_2^2} \times \left\{\mathrm{Im}\left[e^{-i\psi}(A_1 n_1 + A_2 n_2)\right]\right\}^2 + \ldots \tag{28}$$

If we let

$$e^{-i\psi} n_1 \equiv u_1 + iv_1$$

$$e^{-i\psi} n_2 \equiv u_2 + iv_2 \quad ,$$

then

$$\begin{aligned} P - Q &= u_1^2 + v_1^2 + u_2^2 + v_2^2 - u_1^2 + v_1^2 - u_2^2 + v_2^2 \\ &\quad - \frac{2}{A_1^2 + A_2^2}(A_1 v_1 + A_2 v_2)^2 + \ldots \\ &= 2(v_1^2 + v_2^2) - 2\,\frac{(A_1 v_1 + A_2 v_2)^2}{A_1^2 + A_2^2} + \ldots \\ &= 2\,\frac{(A_1 v_2 - A_2 v_1)^2}{A_1^2 + A_2^2} + \ldots \end{aligned}$$

To this order, we put

$$P - Q = \xi^2 \tag{29}$$

where

$$\xi = \left(\frac{2}{A_1^2 + A_2^2}\right)^{\frac{1}{2}} (A_1 v_2 - A_2 v_1) \tag{30}$$

Since v_1 and v_2 are independent Gaussian variables with mean zero and variance σ_a^2, ξ is Gaussian, with mean zero, and variance

$$E\,\xi^2 = \frac{2}{A_1^2 + A_2^2} E\,(A_1 v_2 - A_2 v_1)^2 = 2\,\sigma_a^2 \tag{31}$$

Returning to E (3 + 5/2), we note that the union of regions D_3 and D_5 is contained in the set

$$P - Q \geq \frac{2R}{R-1} (\mu_5 - \mu_2) ,$$

which lies below line d in Figure 1. Thus

$$\begin{aligned} E\,(3 + 5/2) &< \text{Prob} \left\{ \xi^2 \geq \frac{2R}{R-1} (\mu_5 - \mu_2) \right\} \\ &= \text{Prob} \left\{ |\xi| \geq \left[\frac{2R}{R-1} (\mu_5 - \mu_2) \right]^{\frac{1}{2}} \right\} \end{aligned} \tag{32}$$

which is a simple error function. It is a remarkable fact that this error probability is insensitive to signal - to - noise ratio (provided the latter is large).

If hypothesis 5 is true, and the signal power is large compared to the total of receiver noise and interference power, then equations (29) through (31) remain valid, with σ_a^2 replaced by $\sigma_a^2 + \sigma_b^2$, hence an error function just like (32) expresses the approximate probability of declaring H_3 or H_5 when H_5 is true; that is the probability of correctly recognizing the presence of side/back lobe interference along with a signal.

We have left some loose ends in this problem, but our intention was to illustrate the general method of multiple - hypotheses testing, which appears to have some practical utility.

REFERENCES

1. E.M. Hofstetter and D.F. DeLong, "Detection and Parameter Estimation in an Amplitude - Comparison Monopulse Radar", IEEE Transactions on Information Theory, Vol. IT-15, No. 1, pp 22-30, January 1969.

2. Private Communication.

3. Private Communication.

4. Private Communication.